Big data for road safety
– For Safety –

Jim Teressonok
Transportation Engineer, Research Assistant

Together with:
Katerina Folla, Apostolos Ziakopoulos, Dimitris Nikolaou and George Yannis
Background

Contemporary road safety analyses still rely too much on traditional data collection methodologies, so new ways of extracting data are necessary in the new technological landscape.

Big Data might be the solution that will assist researchers, practitioners and decision makers:
- Real-time aspect of data
- Easy acquisition from various sources
  - Smartphones, OBD, cellular networks etc.
- Easy transfer and storage of data to other projects and/or for later use
- Time and cost efficient
Research Questions

- Exploration of the availability of big data that could provide additional insight on road safety analysis.

- Examination of existing data sources with high-resolution data useful for road safety analyses at both macroscopic and microscopic level.

- Exploration of available data useful for a multi-layer approach of road safety analysis:
  - Risk exposure data
  - Road safety performance indicators
  - Accident data
Methodological Challenges

- **Investigation** of all available sources of data
  - Navigation companies were examined
  - More specialized solutions from other companies

- **Reorganisation and categorisation** of the provided data from the companies
  - Different meanings and formats in each company

- **Developing a common analysis structure**
  - Complex data structures
  - Universal description
  - Easy to read and analyse
Risk Exposure Data

- **Road Network** (Road Type, Area Type)
  - High Availability: Google, HERE, TomTom

- **Roadway Geometry** (Horizontal & Vertical Curvature)
  - High Availability: Google, HERE, TomTom

- **Vehicles** (Vehicle Type, Engine Size, Fuel Type)
  - Low Availability: Waze, Uber, IdealSpot

- **Driver Characteristics** (Age, Gender, Nationality)
  - Low Availability: Streetlight (mainly demographics)

- **Distance Parameters** (Person- & Vehicle-kms of Travel)
  - Low Availability: INRIX, Streetlight

- **Traffic Parameters** (Traffic Flow, Density, Volume)
  - High Availability: INRIX, Streetlight, IdealSpot
Road Safety Performance Indicators

➢ **Speed** (Mean Speed, Temporal Variation, Speed Limit)
  • High Availability: HERE, TomTom, INRIX

➢ **Road User** (Seatbelt & Helmet Wearing Rate)
  • No Availability: Alternative Sources

➢ **Traffic Law Enforcement** (Controls, Speed Cameras)
  • Low Availability: Waze, HERE, TomTom, Yandex

➢ **Roads** (Proportion of travel by type of road)
  • Low Availability: INRIX, Streetlight

➢ **Post-Crash Care** (Mean EMS Response Time)
  • No Availability: Alternative Sources
Accident Data

The crash-related data comes in a uniform format:

• Number of Incidents (Including Crashes)
• Type of Delay
• Start & End Location of Crashes
• Road Name
• Type of Area/Road
• Length (In Time) of the Delay
• Weather Conditions
• Significance

Primary Sources: Waze, HERE, TomTom, Yandex, INRIX

Data Extraction Methods: Crowdsourcing, Partnerships, Algorithmically generated flow-based incidents
Alternative Data Sources

- **Logistics and Fleet Operators**
  - UPS, DHL, XPO Logistics, etc.

- **Automotive Manufacturers**
  - Tesla, VW, Toyota, BMW, Ford, etc.

- **Telecommunication Agencies**
  - Verizon, AT&T, Vodafone Group, etc.

- **Carsharing Companies**
  - Uber, Lyft, Cabify, Careem, Easy Taxi, etc.

- **Social Media Platforms**
  - Facebook, Twitter, VK, Instagram, etc.
Scientific and Social Impact

- Alternative data that could lead to **new road safety analyses** in order to:
  - more efficiently describe the road safety phenomenon
  - be used for the validation of traditional research results

- New increased **net present value of road safety data**, available for early problem detection and prompt and customised decision support

- **Continuous driver support** with aim to improve driver behavior and develop better road safety culture at all road users, stakeholders and the Authorities
Future Challenges

- Adapting to **new environments** and ways of extracting data in the future
- Exploring **alternative sources of data** concerning variables with limited information
- **Enrichment and improvement of existing data** from users themselves that will contribute to the system in order to improve road safety
- Inspection of **new statistical methods** in order to increase the accuracy of road safety analysis results by using Big Data
- Development and continuous update of policies and processes in order to ensure **personal data privacy and security**.
Big data for road safety
– For Safety –

Jim Teressonok
Transportation Engineer, Research Assistant

Together with:
Katerina Folla, Apostolos Ziakopoulou, Dimitris Nikolaou and George Yannis